

# TeleChem International, Inc. // arrayit.com

Manufacturers of ArrayIt™ Brand Products and Distributors of Fine Chemicals

## e-library

[e-library home](#)

[Alphabetical Listing](#)

[Home Page](#)

[Products](#)

[Services](#)

### Citation

Pe'er, I., Shamir, R. Spectrum alignment: efficient resequencing by hybridization. *Conf Intell Syst Mol Biol* 8:260-268, 2000.

### Summary

Proc Int Conf Intell Syst Mol Biol 2000;8:260-8

Spectrum alignment: efficient resequencing by hybridization.

Pe'er I, Shamir R.

Department of Computer Science, Tel Aviv University, Israel. [izik@math.tau.ac.il](mailto:izik@math.tau.ac.il)

Recent high-density microarray technologies allow, in principle, the determination of the k-mers that appear along a DNA sequence, for  $k = 8 - 10$  in a single experiment. The k-mer contents, also called the spectrum of the sequence, are sufficient to uniquely reconstruct a sequence longer than a few hundred bases. We have devised a polynomial algorithm that reconstructs the sequence, given the spectrum of a homologous sequence. This situation occurs, for example, in the identification of nucleotide polymorphisms (SNPs), and whenever a homologue of the target sequence is known. The algorithm is robust, can handle errors in the spectrum and assumes knowledge of the k-mer multiplicities. Our simulations show that with realistic numbers of SNPs, the algorithm correctly reconstructs a target sequence of length up to 1000 nucleotides when a polymorphic sequence is known. The technique is general and can handle profiles and HMMs as input instead of a single homologous sequence.

### Electronic Manuscript

None available



## Technical Program

# ISMB 2000

[ISMB](#) | [Agenda](#) | For additional assistance, please consult the [Contact Information](#) page.

---

**August 19, Saturday**

**Tutorials - All Day**

**August 20, Sunday**

[August 21 \(Mon\)](#) | [August 22 \(Tues\)](#) | [August 23 \(Weds\)](#)

---

### Networks and Modeling

**Moderator: TBA**

**Keynote**

**08:30 - 09:15**

Gerald Edelman, The Neurosciences Institute

*How Matter Becomes Imagination: From Brain Dynamics to Consciousness*

**09:15 - 09:40**

Eugene P. van Someren, Lodewyk F.A. Wessels, Marcel J.T. Reinders, Delft University of Technology

[Linear Modeling of Genetic Networks from Experimental Data](#)

**09:40 - 10:05**

Mathias Fellenberg, Biomax Informatics GmbH

Max-Planck-Institute f. Biochemistry

Kaj Albermann, Alfred Zollner, Biomax Informatics GmbH

H.W. Mewes, Max-Planck-Institute f. Biochemistry

Jean Hani, Biomax Informatics GmbH

[Integrative Analysis of Protein Interaction Data](#)

**Coffee**

**10:05 - 10:30**

**10:30 - 10:55**

Maura Cárdenas-García, Jaime Lagúnez-Otero, Instituto de Química, UNAM, Ciudad Universitaria

Nikolai Korneev, Instituto Nacional de Astrofísica Óptica y Electrónica

[Efficient Attractor Analysis Based on Self-dependent Subsets of Elements--An Application to Signal Transduction Studies](#)

**10:55 - 11:20**

Yukako Tohsato, Hideo Matsuda, Akihiro Hashimoto, Osaka University

[A Multiple Alignment Algorithm for Metabolic Pathway Analysis Using Enzyme Hierarchy](#)

**11:20 - 11:45**

Michael P.S. Brown, HNC

[Small Subunit Ribosomal RNA Modeling Using Stochastic Context-Free Grammars](#)

**11:45- 01:30****Lunch, Job Fair Recruitment Event, and BOFs****Networks and Modeling Related to Microarrays****Moderator: TBA****Keynote****01:30 - 02:15**

J. Andrew McCammon, Chemistry and Biochemistry, UC San Diego

*Dynamics of Molecular Recognition***02:15 - 02:40**

Soumya Raychaudhuri, Joshua M. Stuart, Xuemin Liu, Peter M. Small, Russ B. Altman, Stanford University

*Pattern Recognition of Genomic Features with Microarrays: Site Typing of Mycobacterium**Tuberculosis**Strains***02:40 - 03:05**

Andrea Califano, Gustavo Stolovitzky, Yuhai Tu, IBM Computational Biology Center

*Analysis of Gene Expression Microarrays for Phenotype Classification***Coffee****03:05 - 03:30****03:30 - 03:55**

Roded Sharan, Ron Shamir, Tel-Aviv University

*CLICK: A Clustering Algorithm with Applications to Gene Expression Analysis***03:55 - 04:20**

Yizong Cheng, University of Cincinnati

George M. Church, Harvard Medical School

*Biclustering of Expression Data***04:20 - 04:45**

Norbert Brändle, Horng-Yang Chen, Horst Bischof, Vienna University of Technology; Hilmar Lapp

Novartis Research Institute, Vienna

*Robust Parametric and Semi-parametric Spot Fitting for Spot Array Images***05:00 - 07:00** Posters, Software Demos, Exhibits**07:00 - 09:00** Industry Night**August 21, Monday**[Back to the Top](#) | [August 20 \(Sun\)](#) | [August 22 \(Tues\)](#) | [August 23 \(Weds\)](#)**Protein Structure and Modeling I**

**Moderator: TBA****Keynote****08:30 - 09:15**

Harold A. Scheraga, Chemistry and Chemical Biology, Cornell University

*Ab Initio Folding of Proteins***09:15 - 09:40**

Piero Fariselli, Rita Casadio, University of Bologna

*Prediction of the Number of Residue Contacts in Proteins***09:40 - 10:05**

Iddo Friedberg, Tommy Kaplan, Hanah Margalit, The Hebrew University-Hadassah Medical School

*Glimmers in the Midnight Zone: Characterization of Aligned Identical Residues in Sequence-Dissimilar Proteins Sharing a Common Fold***Coffee****10:05 - 10:30****10:30 - 10:55**

Pierre Baldi, Gianluca Pollastri, University of California Irvine; Claus A. F. Andersen, Søren Brunak, The Technical University of Denmark

*Matching Protein  $\beta$ -Sheet Partners by Feedforward and Recurrent Neural Networks***10:55 - 11:20**

Maxim Shatsky, Zipora Y. Fligelman, Tel Aviv University, Ruth Nussinov, Tel Aviv University, IRSP-SAIC Lab of Experimental and Computational Biology

Haim J. Wolfson, Tel Aviv University

*Alignment of Flexible Protein Structures***11:20 - 11:45**

Lukasz Jaroszewski, Adam Godzik, The Burnham Institute

*Search for a New Description of Protein Topology and Local Structure***11:45 - 12:10**

J. Michael Sauder; Roland L. Dunbrack Jr., Fox Chase Cancer Center

*Genomic Fold Assignment and Rational Modeling of Proteins of Biological Interest***Lunch and BOFs****12:10 - 01:30****Protein Structure and Modeling II****Moderator: TBA****Keynote****01:30 - 02:15**

Leroy Hood, Institute for Systems Biology, University of Washington

*Computing Life and Biological Complexity*

02:15 - 02:40

Golan Yona, Michael Levitt, Stanford University

Towards a Complete Map of the Protein Space Based on a Unified Sequence and Structure Analysis of All Known Proteins

### Gene Regulation I

Moderator: TBA

02:40 - 03:05

Harmen J. Bussemaker, University of Amsterdam

Hao Li, University of California, San Francisco

Eric D. Siggia, Rockefeller University

Regulatory Element Detection Using a Probabilistic Segmentation Model

Coffee

03:05 - 03:30

03:30 - 03:55

Jaak Vilo, Alvis Brazma, European Bioinformatics Institute EBI

Inge Jonassen, University of Bergen

Alan Robinson, European Bioinformatics Institute EBI

Esko Ukkonen, University of Helsinki

Mining for Putative Regulatory Elements in the Yeast Genome using Gene Expression Data

03:55 - 04:20

Steve Hampson, Pierre Baldi, Dennis Kibler, Suzanne B. Sandmeyer, University of California Irvine

Analysis of Yeast's ORF Upstream Regions by Parallel Processing, Microarrays, and Computational Methods

04:20 - 04:45

Ian Holmes, University of California Berkeley

William J. Bruno, Los Alamos National Laboratory

Finding Regulatory Elements Using Joint Likelihoods for Sequence and Expression Profile Data

05:00 - 07:00 Posters, Software Demos, Exhibits

## August 22, Tuesday

[Back to the Top](#) | [August 20 \(Sun\)](#) | [August 21 \(Mon\)](#) | [August 23 \(Weds\)](#)

---

### Sequence - Alignments, Families and Phylogeny I

Moderator: TBA

Keynote

08:30 - 09:15

Gene Myers, Celera Genomics Corporation, Rockville, MD

***A Whole Genome Assembly of *Drosophila* and a Progress Report on the Human Genome***

**09:15 - 09:40**

Eleazar Eskin, William Noble, Columbia University

Yoram Singer, Hebrew University

***Protein Family Classification using Sparse Markov Transducers***

**09:40 - 10:05**

Itsik Pe'er, Ron Shamir, Tel Aviv University

**Spectrum Alignment: Efficient Resequencing by Hybridization**

**Coffee**

**10:05 - 10:30**

**10:30 - 10:55**

Bogdan Dorohonceanu, C.G. Nevill-Manning, Rutgers University

**Accelerating Protein Classification Using Suffix Trees**

**10:55 - 11:20**

Rainer Spang, Duke University; Marc Rehmsmeier, Jens Stoye, German Cancer Research Center

**Sequence Database Search Using Jumping Alignments**

**11:20 - 11:45**

Mary E. Cosner, Ohio State University

Robert K. Jansen, University of Texas-Austin

Bernard M.E. Moret, University of New Mexico

Linda A. Raubeson, Central Washington University

Li-San Wang, Tandy Warnow, Stacia Wyman, University of Texas-Austin

**A New Fast Heuristic for Computing the Breakpoint Phylogeny and Experimental Phylogenetic Analyses of Real and Synthetic Data**

**11:45 - 12:10**

Dan Gusfield, University of California, Davis

**A Practical Algorithm for Optimal Inference of Haplotypes from Diploid Populations**

**Lunch, BOFs, and ISCB Business Session (open meeting)**

**12:10 - 01:50**

**Gene Regulation II**

**Moderator: TBA**

**01:50 - 02:15**

Alexander Zien, Robert Kuffner, Ralf Zimmer, Thomas Lengauer  
GMD-German National Research Center for Information Technology  
**Analysis of Gene Expression Data with Pathway Scores**

**02:15 - 02:40**

Pavel A. Pevzner, Sing-Hoi Sze, University of Southern California  
**Combinatorial Approaches to Finding Subtle Signals in DNA Sequences**

**02:40 - 03:05**

Saurabh Sinha, Martin Tompa, University of Washington  
**A Statistical Method for Finding Transcription Factor Binding Sites**

**Coffee****03:05 - 03:30****03:30 - 03:55**

Vineet Bafna, Daniel H. Huson, Celera Genomics Corp.  
**The Conserved Exon Method for Gene Finding**

**04:00 - 06:00** Posters, Software Demos, Exhibits**06:00 - 10:00** ISMB 2000 Reception, San Diego Zoo**August 23, Wednesday**

[Back to the Top](#) | [August 20 \(Sun\)](#) | [August 21 \(Mon\)](#) | [August 22 \(Tues\)](#)

---

**Methods****Moderator: TBA****08:30 - 09:15**

**Keynote: G. Christian Overton Lecture: Reading the Book of Life**  
David B. Searls  
SmithKline Beecham Pharmaceuticals

**09:15 - 09:40**

Hagit Shatkay, Stephen Edwards, W. John Wilbur, Mark Boguski, National Center for Biotechnology Information

**Genes, Themes, and Microarrays Using Information Retrieval for Large-Scale Gene Analysis**

**09:40 - 10:05**

Chris Bailey-Kellogg, John J. Kelley III, Cliff Stein, Bruce Randall Donald, Dartmouth College  
**Reducing Mass Degeneracy in SAR by MS by Stable Isotopic Labeling**

**Coffee****10:05 - 10:30**

10:30 - 10:55

Vanathi Gopalakrishnan, Bruce G. Buchanan, John M. Rosenberg, University of Pittsburgh  
*Intelligent Aids for Parallel Experiment Planning and Macromolecular Crystallization*

10:55 - 11:20

Denys Proux, Xerox Research Centre Europe  
François Rechenmann, INRIA Rhône-Alpes  
Laurent Julliard, Xerox Research Centre Europe

*A Pragmatic Information Extraction Strategy for gathering Data on Genetic Interactions*

11:20 - 11:45

Robert F. Murphy, Michael V. Boland, Meel Velliste, Carnegie Mellon University  
*Toward a Systematics for Protein Subcellular Location: Quantitative Description of Protein Localization Patterns and Automated Analysis of Fluorescence Microscope Images*

11:45 - 12:10

Zhengyan Kan, Washington University  
Warren Gish, Washington University School of Medicine  
Eric Rouchka, Washington University  
Jarret Glasscock, Washington University School of Medicine  
David States, Washington University

*UTR Reconstruction and Analysis Using Genomically Aligned EST Sequences*

Lunch and BOFs

12:10 - 01:30

**Sequence - Alignments, Families and Phylogeny, II**

**Moderator: TBA**

**Keynote**

01:30 - 02:15

Minoru Kanehisa, Institute for Chemical Research, Kyoto University  
*Graph Comparison and Path Computation Methods for Predicting Molecular Networks from Genome Information*

02:15 - 02:40

Robin McEntire, SmithKline Beecham Pharmaceuticals  
Peter Karp, SRI-International  
Neil Abernethy, InGenuity  
David Benton, SmithKline Beecham Pharmaceuticals  
Gregg Helt, University of California Berkeley  
Matt DeJongh, NetGenics  
Robert Kent, Ontologos  
Anthony Kosky, GeneLogic  
Suzanna Lewis, University of California Berkeley  
Dan Hodnett, NetGenics  
Eric Neumann, 3rd Millenium  
Frank Olken, Lawrence Berkeley Livermore Laboratory

Dhiraj Pathak, SmithKline Beecham Pharmaceuticals

Luca Toldo, Merck KgaA

Thodoros Topaloglou, GeneLogic

*An Evaluation of Ontology Exchange Languages for Bioinformatics*

**02:40 - 03:05**

Mark Craven, David Page, Jude Shavlik, Joseph Bockhorst, Jeremy Glasner, University of Wisconsin

*A Probabilistic Learning Approach to Whole-Genome Operon Prediction*

**Coffee**

**03:05 - 03:30**

**03:30 - 03:55**

Stefan Kurtz, Enno Ohlebusch, Chris Schleiermacher, University of Bielefeld

Jens Stoye, German Cancer Research Center

Robert Giegerich, University of Bielefeld

*Computation and Visualization of Degenerate Repeats in Complete Genomes*

**03:55 - 04:20**

Mathieu Blanchette, Benno Schwikowski, Martin Tompa, University of Washington

*An Exact Algorithm to Identify Motifs in Orthologous Sequences from Multiple Species*

**04:20 - 04:45**

Best Poster and Paper Awards Session

---

[Back to the Top](#) | [August 20 \(Sun\)](#) | [August 21 \(Mon\)](#) | [August 22 \(Tues\)](#) | [August 23 \(Weds\)](#)

---

[ISMB](#) | [Agenda](#) | For additional assistance, please consult the [Contact Information](#) page.